ACADEMIC MEDICINE
Journal of the Association of American Medical Colleges

Author: Thomas F. Heston, MD, MS

Title: Statistical Fragility in Surveys

DOI: 10.1097/ACM.0000000000005584

Academic Medicine

DOI: 10.1097/ACM.0000000000005584

Statistical Fragility in Surveys

To the Editor: I read with interest the article by Carnes and colleagues examining the effect of a workshop on overcoming bias and improving work climate in departments of medicine.[1] These are laudable goals. However, the statistical fragility requires a cautious interpretation of the results.

The primary issues are the low and uneven response rates (RR), small effect sizes, and the performance of multiple significance tests. Although the RR was high for a physician survey, a rate below 50% still increases potential nonresponder bias. Weighting analyses may help, but they also add more assumptions. Additionally, slight differences in RR have a disproportionate impact when the sample size is large, and the RR is low. In this survey, the RR between groups differed significantly before and after the workshop. Also, the decline in RR was more significant in the intervention group compared to controls (6.3% versus 3.9%). This difference could have easily changed the statistical analysis.

Secondly, effect sizes were small. For example, there was less than a 0.1 change on the 5-point scale for climate. While statistically significant, the cause could have been from any number of influences (e.g., a different time of year or day of the week). Attributing these slight differences solely to a 3-hour workshop stretches credulity.

Finally, performing multiple significance tests increases the risk of type I errors. If correcting for the false discovery rate was applied, no comparisons would have been statistically significant.[2]

Perhaps the most interesting finding is that nearly half of the respondents self-identified as a minority (47%). Although respondents were asked if they had "any other self-identified minority status," specifying additional minority categories may have influenced the results to the point where the majority self-identified as a minority.

Although this survey had fragile results, the researchers should be commended for undertaking an intervention aimed at the admirable goal of reducing bias and improving workplace climate. Going forward, seeking higher response rates and larger effect sizes will help demonstrate whether brief interventions can lead to meaningful improvements on these critical issues.

Funding/Support: None reported.

Other disclosures: None reported.

Ethical approval: Reported as not applicable.

Thomas F. Heston, MD, MS

Clinical associate professor, Medical Education and Clinical Sciences, Washington State University – Spokane, Spokane, Washington; email: tom.heston@wsu.edu

First published online

References

1. Carnes M, Sheridan J, Fine E, Lee Y-G, Filut A. Effect of a workshop to break the bias habit for internal medicine faculty: A multisite cluster randomized controlled study. Acad Med. 2023;98:1211-1219.
2. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol. 1995;57:289-300.